Viral Evolution and Adaptation as a Multivariate Branching Process

Fernando Antoneli,[1,*] Francisco Bosco,[2,3,†] Diogo Castro,[4,‡] and Luiz Mario Janini[3,5,§]

[1] *Departamento de Informática em Saúde, Universidade Federal de São Paulo, São Paulo, SP, Brazil.*
[2] *Laboratório de Biocomplexidade e Genômica Evolutiva, Universidade Federal de São Paulo, São Paulo, SP, Brazil.*
[3] *Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de São Paulo, São Paulo, SP, Brazil.*
[4] *Departamento de Infectologia, Universidade Federal de São Paulo, São Paulo, SP, Brazil.*
[5] *Departamento de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil.*

(Dated: December 21, 2011)

In the present paper we analyze the problem of adaptation and evolution of RNA virus populations, by defining the basic stochastic model as a multivariate branching process. The defined stochastic process turns out to be well suited to describe several aspects of RNA viral populations. We show that in the absence of beneficial forces the model is exactly solvable. As a result it is possible to prove several key results directly related to known typical properties of these systems. Moreover, new insights on the dynamics of evolving virus populations can be foreseen.

PACS numbers: 87.10.Mn, 87.23.Kg, 87.23.Cc ,87.10.Ca
Keywords: Viral Evolution, Branching Process

I. INTRODUCTION

RNA viruses exhibit a pronounced genetic diversity [1]. This variability allows RNA virus to better adapt to environmental challenges as represented by host and therapy pressures [2]. Due to the lack of a proofreading activity of viral RNA polymerases (average error incorporation rate in the order of 10^{-4} per nucleotide, per replication cycle [3]), short generation times and huge population numbers, RNA viral populations may be viewed as a collection of particles bearing mutant genomes. As a consequence of high mutation rates, frequencies of mutants depend not only on their level of adaptation but on the probability of being faithfully replicated during viral genomic RNA synthesis. Several studies have looked at viral diversification processes as a contributing cause of disease progression and of therapeutic strategies shortcomings including vaccine trials [2, 4]. It has become important to understand the process by which virus acquire diversity and the dynamics and fluctuations of this diversity in time. However, understanding viral evolution *in vivo* has proven to be a very cumbersome accomplishment due to the so many variables present in the interplay between virus and their hosts. To name a few; the host defense pressures as the innate and "cognitive" immune responses, the use of antiviral drugs, the turnover rate of virus populations composed by viral replication and clearance, the elevated mutational rates of RNA virus, and the possible existence of structured viral reservoirs in infected patients. It is also important to take into account the size of the viral innoculums at the moment of infection, how frequent viral populations undergo bottleneck passages within a host, how differently each infected individual may react to an incoming virus and finally how many viral variants are tightly associated with differential biological capabilities.

Traditionally, in an effort to make the viral evolution process more palpable, several groups have addressed this subject from different points of view. There is a substantial amount of publications that studied virus populations during their evolution in experimental settings, for instance, cell cultures [5], by challenging the virus with population bottlenecks [6, 7], or the introduction of antiviral drugs [8], including mutagens, or another competing viral population. Experiment outcomes were evaluated using viral replication kinetics, the intensity and quality of the observed mutational spectra and virus survival/extinction as final parameters. Several other groups have studied the process of viral evolution away from the bench but using mathematical and computational tools [9–14]. These models are quite tractable but there is always the risk of oversimplification. To escape from oversimplifying the interplay between virus and hosts a model needs to incorporate a few hard rules based on previous experimental data which has been generated by the whole community of investigators addressing viral evolution. Based on other groups experimental data and previous mathematical models put forward by other investigators as the one presented by Lázaro *et al.* [10] we sought to study a stochastic model for virus evolution that would be able to describe some general aspects of RNA virus evolution. Here, RNA viral evolution is described by a multivariate branching process during which each round of replication is accompanied by the introduction of a single point mutation per genome in the viral progeny. Drake and Holland [15] back in 1999 have inferred, based on limited data, a central value for the RNA virus mutation rate per genome per replication

* fernando.antoneli@unifesp.br; corresponding author.
† fbosco@unifesp.br
‡ diogo.castro@unifesp.br
§ janini@unifesp.br

of $\mu_g \approx 0.76$ and suggested the rate per round of cell infection of $\mu_g \approx 1.5$. In 2010, Sanjuán *et al.* [16] revisiting this theme by reviewing a list of previous publications encountered RNA virus mutational rates in the order of 10^{-4} to 10^{-6} with $\mu_g \approx 4.64$ for the bacteriophage Qβ (Batschelet *et al.* [17]) and $\mu_g \approx 1.15$ for hepatitis C virus (Cuevas *et al.* [18]).

It has been demonstrated that virus populations may be reduced at the moment of infection, and only a few particles are able to start a new infection process in naive hosts [19, 20]. Abrupt reductions on RNA viral populations known as population bottlenecks may eliminate population diversity and lead the virus to pathways towards extinction due to the exacerbated effects of genetic drift. An incoming virus population recovering from a transmission bottleneck event may show an asymptotic behavior resembling stationary equilibrium represented by the balance between two opposite forces classically identified with mutation and selection. This asymptotic behavior would occur if the environment is constant and enough time is allowed between two successive bottleneck events. The relaxation time between the bottleneck and the establishment of stationary equilibrium has been referred to as the "recovering time" by Aguirre *et al.* [14].

It has been pointed out (Drake and Holland [15]) that the basal value of RNA virus mutation rates is so large and RNA virus genomes are so informationally dense, that even a modest rate increase extinguishes the population. The frequent appearance of overlapping reading frames and multifunctional proteins augments the risk of a random mutation to have a deleterious impact and even more, multiply the effect of deleterious mutations. For example, the fraction of deleterious mutations out of random mutations occurring in vesicular stomatitis virus is around 70% (Sanjuán *et al.* [21]). If the introduction of a mutagen to a replicating virus population is able to cause its extinction by increasing mutational rates, the process is known as chemical lethal mutagenesis and has been demonstrated in a number of viruses including the vesicular stomatitis virus (VSV) [22, 23], human immunodeficiency virus type 1 (HIV-1) [24], poliovirus type 1 [8, 22], foot-and-mouth disease virus [25], lymphocytic choriomeningitis virus [26], Hanta virus [27] and Hepatitis C virus [28]. Accordingly, in the model, increases on mutational rates, and more specifically, on the deleterious component of the mutational spectrum are able to push viral populations towards extinction. Our results corroborate with the study from Bull, Sanjuán and Wilke [12] by showing that the sufficient condition for lethal mutagenesis involves mutational and ecological aspects. Bull *et al.* [12] arrived at a conjectural criteria for lethal mutagenesis by a heuristic and intuitive approach of possible general applicability. By applying the branching process theory to the evolution of RNA viruses the lethal mutagenesis inequality proposed by Bull *et al.* [12] is rigorously proven here. Furthermore, we describe four distinct regimes of RNA virus populations: transient regime, stationary equilibrium, extinction threshold, and extinction through lethal mutagenesis.

The approach we adopt to the problem of virus adaptation and evolution allows us to prove several key results directly related to known typical properties of these systems and to get new insights on the dynamics of evolving virus populations.

We note that in previous works [11, 14, 29–31] the properties of phenotypic models are discussed starting from a mean field linear model described by a mean matrix without reference to the underlying stochastic process modeling the microscopic dynamics of particle replication. Here we deduce the matrix of first moments from a generating function of the stochastic process. Although the two matrices happen to coincide, it is important to stress that only from the generating function of the underlying stochastic process it is possible to fully discuss the validity of the model.

Structure of the Paper. In section II we describe a class of models for viral evolution and show that they define a multitype branching process. We explicitly compute the generating function and derive some elementary properties. In section III we briefly recall some basic concepts and results from the theory of multitype branching processes. In section IV we solve the spectral problem for the mean matrix of the model which allows us to apply the results from section III. Finally, in section V we present our conclusions and directions for future research.

II. PHENOTYPIC MODELS FOR VIRAL EVOLUTION

In this section we describe a model for viral evolution and will show that it is naturally represented by a multivariate branching stochastic process. Our description is along the same line of the probabilistic model introduced by Lázaro *et al.* [10]. We interpret the notion of mutation probability as a general effect of probabilistic nature acting on the replication capability of individual viral particles, considered here as a measure of the particle's fitness characterizing its phenotype. This effect is summarized by the definition of a stationary probability distribution which is used to set up a Galton-Watson branching process (Watson and Galton [32]) for the temporal evolution of the viral population. This probability distribution gives appropriate parameters to classify the asymptotic behavior of the viral population and to describe some of the non-equilibrium properties of the model.

In other related publications the concept of mutation is extensively used as the cause of replication capacity change. Understanding that those changes constitute an observable output due to many different factors (of genetic and non-genetic nature), we prefer to use the general term "effect" over the replication capacity to characterize the three possible changes (deleterious, beneficial and neutral) that may happen with the viral particle

when it replicates. The precise definition of the three types of changes are given in the next section.

A. Definition of the Model

A number of viral infections starts with the transmission of a relatively small number of viral particles from one organism to another one. The initial viral population starts replicating constrained by the unavoidable interaction with the host organism and evolves in time towards an eventual equilibrium. Each particle composing the population replicates in the cellular context that may differ from cell to cell. Moreover each particle has different replication capabilities due to the natural genomic diversity found in viral populations in general. Therefore, it is reasonable to consider the viral population as a set of particles divided in groups of different replication capabilities measured in terms of the number of particles that one particle can produce. Each of those groups we call a class; the replication capability of a viral particle is an output of the process of interaction of that particle carrying its genetic information with the cell environment. The replication capability is considered as a phenotypic character of the particle and therefore each class is considered as a set of particles with a possible genotype diversity expressing the same phenotype trait. The model we consider here does not take into account any information about the genomic diversity of any replicating class and therefore it should be classified as phenotypic model.

We consider that the whole set of particles composing the viral population replicates at the same time in such a way that the evolution of the population is described as a succession of discrete viral generations. This assumption crucially depends on the clear definition of the time needed for a particle to replicate, referred by virologists as *generation time*. As it depends on the cell environment it is clear that this time period may vary from particle to particle replicating in different cells in such a way that the meaningful concept is a distribution of replication times with a possible clear mean value. The dispersion of the replication times can be considered small if we restrict ourselves to homogeneous cell populations. Under these conditions we consider that no particle can be part of two successive generations. The possible impact of a subset of non replicating particles on the dynamics of the viral population is left to further studies.

Suppose that we have a population of viruses that start evolving from an initial set of particles (population at $t = 0$), which is partitioned into *classes* according to the *replication capacity* of each particle, that is, where each particle of class 0 produces no copies of itself, each particle with class 1 produces one copy of itself, and so on. We assume that there is a *maximum replication capacity* R imposed by the natural limiting conditions under which any particle of the population replicates. Moreover, as the process of replication is controlled by chemical reactions involving specific enzymes and the template, it is reasonable to assume a mean bounded replication capacity per particle that is possibly typical for each specific virus.

In the process of replication of a viral particle errors may occur at each replication cycle in the form of point mutations with possible impact on the replication capacity of the progeny particles. Due to the intrinsic stochastic component of chemical reactions it is natural to treat this point mutational cause as probabilistic. Another possible cause of change in the replication capability in the viral offspring is clearly related to the cellular environment where the replication process takes place. As a result the time evolution of viral populations should be viewed as a physical process strongly influenced by stochasticity. Therefore we consider that the combined action of genetic and non-genetic causes may produce basically three types of replicative effects namely: *deleterious* with decreased replication capability from the parental to the progeny particles, *beneficial* when the replication capability increases and *neutral* with no change of the particles' replication capacity. In the model under consideration we define these three effects as:

- *deleterious effect*: the replication capacity of the copied particle decreases by one. Note that when the particle has capacity of replication equal to 0 it will not produce any copy of itself.
- *beneficial effect*: the replication capacity of the copy increases by one. If the replication capacity is already the maximum allowed then the replication capacity of the copies will stay the same.
- *neutral effect*: the replication capacity of the copies remain the same as the replication capacity of the original parental particle.

For each type of effect we associate a probability at the particle scale applicable to every single replication event: d for the probability of the occurrence of a *deleterious* effect in one replication cycle of each particle, b for the probability of the occurrence of a *beneficial* effect in one replication cycle of each particle. The complementary probability $c = 1 - b - d$ is the probability of the occurrence of a neutral effect. In the case of *in vitro* experiments with homogeneous cell populations the parameters c, d and b may be considered as mutation probabilities.

The *simple phenotypic model* is obtained by requiring that there are no beneficial effects in time, that is $b = 0$. This assumption is justified by several experimental results. The frequencies between beneficial, deleterious and neutral mutations appearing in a replicating population have been already measured by prior studies [21, 33–40]. Taking their results together, it is reasonable to conclude that beneficial mutations could be as low as 1000 less frequent than either neutral or deleterious mutations. As a result the viral population would be submitted to a large number of successive deleterious and neutral changes and a comparatively small number of beneficial changes.

From what is described above it should become clear that the model assumes a scenario where a probabilistic processes at the cellular/viral scale take place in the context of the interaction between the viral particle and the host cell. The combined effect of small scale processes are observed at the viral population scale in terms of collective (stable or not) properties.

Based on the general aspects of the phenomenon of viral replication it is compelling to to model it in terms of a branching process. In this perspective we define a *discrete multitype Galton-Watson branching process* for the evolution of the initial population, where the *classes* will be represented by the replication capabilities $0, 1, \ldots, R$. The branching process is described by a sequence of vector-valued random variables $\{\boldsymbol{Z}_n : n \in \mathbb{N}\}$ giving the number of particles in each replication class in the n-th generation. Thus $\boldsymbol{Z}_n$ are vectors of non-negative integers satisfying the following assumption: if the size of the n-th generation is known, then the probability laws governing the later generations does not depend on the sizes of generations preceding the n-th, that is the sequence $\{\boldsymbol{Z}_n : n \in \mathbb{N}\}$ forms a *markovian process*. The initial population $\boldsymbol{Z}_0$ is represented by a vector of non-negative integers $\boldsymbol{Z}_0 = (Z_0^0, Z_0^1, \ldots, Z_0^R)$, which is non-zero and non-random. The temporal evolution of the population is obtained from a vector-valued discrete probability distribution $\boldsymbol{\zeta} = (\zeta_0, \zeta_1, \ldots, \zeta_R)$ defined on the set of vectors with non-negative integer entries called the *offspring distribution* of the branching process. For any vector with non-negative entries $\boldsymbol{i} = (i^0, \ldots, i^R)$ one has that

$$\mathbf{P}(\boldsymbol{Z}_{n+1} = \boldsymbol{i} | \boldsymbol{Z}_n = \boldsymbol{e}_r) = \zeta_r(\boldsymbol{i})\,, \tag{1}$$

where $\boldsymbol{e}_r = (0, \ldots, 1, \ldots, 0)$, with 1 in the r-th position. Thus, $\zeta_r(\boldsymbol{i})$ is the joint probability that an individual particle of class r $(0 \leqslant r \leqslant R)$ generates i^0 progeny particles in the class 0, i^1 progeny particles in the class 1, ..., i^R progeny particles in the class R. Note that any vector $\boldsymbol{Z}_n = (Z_n^0, Z_n^1, \ldots, Z_n^R)$ may be written as a sum $\sum_r Z_n^r \boldsymbol{e}_r$ and since each particle in $\boldsymbol{Z}_n$ may be seen as the initial condition of a new branching process independently of the others, equation (1) determines the probability laws for a general branching process as follows

$$\mathbf{P}\big(\boldsymbol{Z}_{n+1} = \boldsymbol{i} | \boldsymbol{Z}_n = \textstyle\sum_r Z_n^r \boldsymbol{e}_r\big) = \displaystyle\prod_r \zeta_r(\boldsymbol{i})^{Z_n^r}\,.$$

In order to compute the offspring probability distribution $\boldsymbol{\zeta}$ for the simple phenotypic model, we start by observing that ζ_r is non-zero only when $\boldsymbol{i}$ is of the form $\boldsymbol{i} = (0, \ldots, i^{r-1}, i^r, \ldots, 0)$ since a particle with replication capability r can only produce progeny particles of the replication capability r or $r-1$, moreover the entries i^{r-1} and i^r should satisfy $i^{r-1} + i^r = r$. Thus we just need to compute the probabilities ζ_r on the vectors of the form $\boldsymbol{i}_k = (0, \ldots, r-k, k, \ldots, 0)$. Suppose that a viral particle v with replication capacity r $(0 \leqslant r \leqslant R)$ replicates itself producing new virus particles $v_1, \ldots, v_r$. For each new particle v_i, there are two possible outcomes regarding the type of change that may occur: neutral or deleterious, with probabilities $c = 1 - d$ and d, respectively. Representing the result of the i-th replication event by a variable X_i, which can assume two values: 0 if the effect is deleterious (failure) and 1 if the effect is neutral (success), the probability distribution of X_i is that of a *Bernoulli trial* with probability of occurrence of a neutral effect $c = 1 - d$ (success), that is,

$$\mathbf{P}(X_i = k) = (1-d)^k\, d^{1-k} \qquad (k = 0, 1)\,.$$

The total number of neutral effects that occur when the original virus particle reproduces is a random variable S_r given by the sum of all variable X_i, since each copy is produced independently of the others,

$$S_r = X_1 + X_2 + \ldots + X_r\,.$$

That is, S_r counts the total number of neutral effects (successes) that occurred in the production of r virus particles $v_1, \ldots, v_r$. It also represents the total number of particles that will have the same replication capacity r of the original particle v. It is well known (Feller [41]) that a sum of r independent and identically distributed Bernoulli random variables with probability $c = 1 - d$ of success has a probability distribution given by the *binomial distribution*:

$$\mathbf{P}(S_r = k) = \mathrm{binom}(k; r, 1-d) = \binom{r}{k} (1-d)^k\, d^{r-k}\,.$$

Since this is the probability that a class r virus particle v produces k progeny particles with the same replication capability as itself one has therefore

$$\zeta_r(0, \ldots, r-k, k, \ldots, 0) = \mathbf{P}(S_r = k) = \mathrm{binom}(k; r, 1-d)\,.$$

Given the offspring probability distribution $\boldsymbol{\zeta}$ one may set up a *probability generating function* $\boldsymbol{f} = (f_0, \ldots, f_R)$ which is defined by the power series

$$f_r(z_0, z_1, \ldots, z_R) = \sum_{\boldsymbol{i}} \zeta_r(\boldsymbol{i})\, z_0^{i^0} \ldots z_R^{i^R}\,.$$

The probability generating function of the simple phenotypic model is

$$\begin{aligned} f_0(z_0, z_1, \ldots, z_R) &= 1 \\ f_1(z_0, z_1, \ldots, z_R) &= dz_0 + cz_1 \\ f_2(z_0, z_1, \ldots, z_R) &= (dz_1 + cz_2)^2 \\ &\vdots \\ f_R(z_0, z_1, \ldots, z_R) &= (dz_{R-1} + cz_R)^R \end{aligned} \tag{2}$$

Note that the functions f_r are polynomials whose coefficients are exactly $\mathrm{binom}(k; r, 1-d)$. This function completely determines the branching process.

Now it is easy to obtain the general case where the beneficial effects have a non-zero contribution b. In this case, the binomial distribution is replaced by a *trinomial*

distribution (see Feller [41]) and the probability generating function of the general phenotypic model is

$$\begin{aligned} f_0(z_0, z_1, \ldots, z_R) &= 1 \\ f_1(z_0, z_1, \ldots, z_R) &= dz_0 + cz_1 + bz_2 \\ f_2(z_0, z_1, \ldots, z_R) &= (dz_1 + cz_2 + bz_3)^2 \\ &\vdots \\ f_{R-1}(z_0, z_1, \ldots, z_R) &= (dz_{R-2} + cz_{R-1} + bz_R)^{R-1} \\ f_R(z_0, z_1, \ldots, z_R) &= (dz_{R-1} + (c+b)z_R)^R \end{aligned} \tag{3}$$

Remark 1 *It is worth to mention other variations of these models that share the same essential properties and are more adequate in different contexts.*

With Zero Class: *In this variation, which is the version deduced above, particles of class $r = 0$ are generated by the particles from class $r = 1$.*

Without Zero Class: *In this variation, the particle class 0 is omitted and thus the probability generating function has R variables and R components: omit the variable z_0, the first component f_0 and define $f_1(z_1, \ldots, z_R) = d + cz_1 + bz_2$. Particles of class $r = 1$ undergoing a deleterious change are eliminated in the next generation.*

B. Basic Properties of the Phenotypic Model

We start by recalling that, when calculating probabilities and expectations, there is no loss of generality if one considers only initial populations consisting of just one particle of class r ($0 \leqslant r \leqslant R$), since the general case can be decomposed as a sum of independent processes with this kind of initial population. All the relevant properties of the model can be deduced with this simplification.

We shall introduce the notation $Z_0^r = 1$ for the condition $\boldsymbol{Z}_0 = \boldsymbol{e}_r$, which is the initial population consisting of one particle of class r and zero particles of other classes. Thus $\mathbf{P}(\boldsymbol{Z}_1 = \boldsymbol{i} | Z_0^r = 1) = \zeta_r(\boldsymbol{i})$. A basic assumption in the theory of branching processes is that all the first moments are finite and that they are not all zero. Then one may consider the *mean evolution matrix* or the *matrix of first moments* $\boldsymbol{M} = \{M_{ij}\}$ which describes how the averages of the sub-populations of particles in each replication class evolves in time:

$$M_{ij} = \mathbf{E}(Z_1^i | Z_0^j = 1)\,, \qquad \forall\, i, j = 0, \ldots, R\,.$$

In terms of the probability generating function one has

$$M_{ij} = \frac{\partial f_j}{\partial z_i}(1, 1, \ldots, 1)\,.$$

Denoting by $\boldsymbol{f}'$ the jacobian matrix of $\boldsymbol{f}$ one may write

$$\boldsymbol{M} = \boldsymbol{f}'(\mathbf{1})\,, \quad \text{where} \quad \mathbf{1} = (1, 1, \ldots, 1)\,.$$

The evolution of the averages $\langle \boldsymbol{Z}_n \rangle$ of $\boldsymbol{Z}_n$ is given by

$$\langle \boldsymbol{Z}_n \rangle = \mathbf{E}(\boldsymbol{Z}_n | \boldsymbol{Z}_0) = \boldsymbol{M}^n\, \boldsymbol{Z}_0\,. \tag{4}$$

From the generating functions (2) and (3) it is trivial to compute the mean matrix of the phenotypic model. For the simple phenotypic model it is

$$\boldsymbol{M} = \begin{pmatrix} 0 & d & 0 & 0 & 0 & \ldots & 0 \\ 0 & c & 2d & 0 & 0 & \ldots & 0 \\ 0 & 0 & 2c & 3d & 0 & \ldots & 0 \\ 0 & 0 & 0 & 3c & 4d & \ldots & 0 \\ 0 & 0 & 0 & 0 & 4c & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & Rd \\ 0 & 0 & 0 & 0 & 0 & 0 & Rc \end{pmatrix} \tag{5}$$

Note that it is an upper triangular matrix. In the case of the general phenotypic model it is

$$\boldsymbol{M} = \begin{pmatrix} 0 & d & 0 & 0 & 0 & \ldots & 0 \\ 0 & c & 2d & 0 & 0 & \ldots & 0 \\ 0 & b & 2c & 3d & 0 & \ldots & 0 \\ 0 & 0 & 2b & 3c & 4d & \ldots & 0 \\ 0 & 0 & 0 & 3b & 4c & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & Rd \\ 0 & 0 & 0 & 0 & 0 & (R-1)b & R(c+b) \end{pmatrix} \tag{6}$$

which is a tri-diagonal matrix. It is clear that d is related to upper part of the matrix, b to the lower part and c to the main diagonal.

The mean matrix of the phenotypic model can be viewed as the adjacency matrix of a directed weighted graph where the nodes represent the particle classes according to their replication capacity and the arrows represent the effect of decrease or increase of the replication capacity due to the replication process (see FIG. 1).

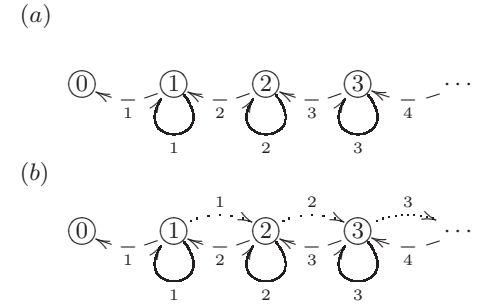

FIG. 1. Graphs of mean matrices. (a) Simple phenotypic model. (b) General phenotypic model. The arrows are numbered according to which there occurs a deleterious effect (d – dashed arrows) or a beneficial effect (b – dotted arrows) or neutral effect (c – solid arrows).

III. RELEVANT RESULTS FROM THE THEORY OF BRANCHING PROCESSES

In this section we collect a few definitions and results from the theory of branching process that will be necessary in our analysis of the phenotypic model.

A. The Mean Matrix of a Branching Process

Consider a multitype branching process $\boldsymbol{Z}_n$ with offspring probability distribution $\boldsymbol{\zeta}$ and probability generating function $\boldsymbol{f}$. Suppose that $\boldsymbol{\zeta}$ has all its first moments finite and not all zero. Then conditioning on the elementary initial populations $Z_0^r = 1$ on may define the *mean matrix* $\boldsymbol{M} = \{M_{ij}\}$ of the multitype branching process $\boldsymbol{Z}_n$ by

$$M_{ij} = \mathbf{E}(Z_1^i | Z_0^j = 1) \quad \forall\, i, j = 0, \ldots, R\,.$$

In general, a multitype Galton-Watson branching process can be classified into *decomposable* and *indecomposable* according to which its mean matrix is reducible or irreducible, respectively. A non-negative matrix $\boldsymbol{M} = \{M_{ij}\}$ $(0 \leqslant i, j \leqslant R)$ is called *irreducible* if for every pair of indices i and j, there exists a natural number n such that $\left(\boldsymbol{M}^n\right)_{ij} > 0$ and it is called *reducible* otherwise (see Gantmatcher [42]). There is another characterization of irreducibility in terms of the graph of the matrix.

The *graph* $\mathcal{G}(\boldsymbol{M})$ of $\boldsymbol{M}$ is defined to be the directed graph on R nodes $\{0, 1, \ldots, R\}$, each corresponding to a type of particle, in which there is a directed edge leading from node i to node j if and only if $M_{ij} \neq 0$. A graph $\mathcal{G}(\boldsymbol{M})$ is called *path connected* if for each pair of nodes (i, j) there is a sequence of directed edges leading from i to j. A matrix $\boldsymbol{M}$ is irreducible if and only if $\mathcal{G}(\boldsymbol{M})$ is path connected (see Meyer [43]).

A multitype Galton-Watson branching process is called *positively regular* if its mean matrix $\boldsymbol{M}$ is *primitive*, that is, $\boldsymbol{M}^n$ is positive for some positive integer n. In particular, a positively regular branching process is indecomposable, since a primitive matrix is irreducible (see Gantmatcher [42] or Meyer [43]). Positive regularity is a standard assumption in the study of multitype branching processes, as it opens up the way to apply the powerful Perron-Frobenius theory (see Harris [44] or Athreya and Ney [45]).

Example 1 *The classification of the phenotypic model according to the irreducibility or reducibility of its mean matrix is the following:*

(i) In the version "with zero class" the mean matrix (6) *or* (5) *will have the first column filled with zeros, that is, they are not primitive matrices and thus the corresponding branching processes are not positively regular. Moreover, a quick look at the graph* $\mathcal{G}(\boldsymbol{M})$ *in FIG. 1 (b) shows that the process is decomposable since the node corresponding to particles of type* 0 *does not have a direct arrow leading to other nodes. In the case of the simple phenotypic model, the corresponding graph* $\mathcal{G}(\boldsymbol{M})$ *is shown FIG. 1 (a). Note that there are no dotted arrows since the probability of beneficial effects is* 0 *and so the graph is* totally path disconnected, *in other words, each "path component"of the graph consists of exactly one node.*

(ii) In the version "without zero class" the mean matrix of both models can be obtained from (6) *and* (5) *by removing the first row and the first column. Now the general phenotypic model becomes positively regular, since the node corresponding to particles of class* 0 *no longer exists. The simple phenotypic model still is decomposable, even without the node corresponding to particles of class* 0.

B. Malthusian Parameter and Extinction Probability

Let $\varrho(\boldsymbol{M})$ denote the *spectral radius* of $\boldsymbol{M}$, that is, if $\lambda_1, \ldots, \lambda_R$ are the eigenvalues of $\boldsymbol{M}$ then

$$\varrho(\boldsymbol{M}) = \max\left\{|\lambda_r|\right\}.$$

Since $\boldsymbol{M}$ is a non-negative matrix, it has at least one largest non-negative eigenvalue which coincides with its spectral radius (see Gantmatcher [42] or Meyer [43]). When the largest eigenvalue is positive we shall call it, following Kimmel and Axelrod [46], the *malthusian parameter* m of the branching process (see also Jagers *et al.* [47]).

The malthusian parameter of a multitype Galton-Watson branching process plays the same role as the mean of the probability distribution of the offspring in a simple Galton-Watson process and its name is motivated by equation (4), which implies that $\varrho(\boldsymbol{M}^n) = m^n$, the average population size increases or decreases at a geometric rate, in accordance with the "Malthusian Law of Growth".

Finally, it follows from the theory of non-negative matrices that there is a *left non-negative eigenvector* $\boldsymbol{v}$ and a *right non-negative eigenvector* $\boldsymbol{u}$ corresponding to the eigenvalue m:

$$\boldsymbol{v}^{\mathrm{t}}\,\boldsymbol{M} = m\,\boldsymbol{v}^{\mathrm{t}} \qquad \text{and} \qquad \boldsymbol{M}\,\boldsymbol{u} = m\,\boldsymbol{u}\,,$$

which can be normalized so that

$$\boldsymbol{v}^{\mathrm{t}}\boldsymbol{u} = 1 \quad \text{and} \quad \mathbf{1}^{\mathrm{t}}\boldsymbol{u} = 1\,, \tag{7}$$

where $\boldsymbol{v}^{\mathrm{t}}$ is the transposed of the vector $\boldsymbol{v}$. Moreover, when $\boldsymbol{M}$ is irreducible the left and right eigenvectors are positive (see Gantmatcher [42] or Meyer [43]).

Let $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_R)$ be the *vector of extinction probabilities*

$$\gamma_r = \mathbf{P}(\boldsymbol{Z}_n = 0 \text{ for some } n | Z_0^r = 1)\,,$$

the probability that the process eventually become extinct given that initially there is exactly one particle of class r. In general, when the initial condition is given by a vector of non-negative integers $\boldsymbol{Z}_0 = (Z_0^0, Z_0^1, \ldots, Z_0^R)$ the extinction probability is

$$\mathbf{P}(\boldsymbol{Z}_n = 0 \text{ for some } n | \boldsymbol{Z}_0) = \prod_{i=0}^{R} {\gamma_i}^{Z_0^i} .$$

A basic result of the theory of branching processes is that the vector of extinction probabilities $\boldsymbol{\gamma}$ is the solution in $[0,1]^R$ with smallest components of the equation

$$\boldsymbol{f}(\boldsymbol{\gamma}) = \boldsymbol{\gamma} , \tag{8}$$

where $\boldsymbol{f}$ is the probability generating function. Observe that $\mathbf{1}$ is always a fixed point of $\boldsymbol{f}$, that is, a solution of equation (8). Therefore, if there is no other solution of equation (8) in the unit cube $[0,1]^R$ then the process always has probability 1 to become extinct.

The main classification result in the indecomposable case, states that there are only three possible regimes (see Harris [44] or Athreya and Ney [45]):

(i) If $m > 1$ then $\mathbf{0} \leqslant \boldsymbol{\gamma} < \mathbf{1}$ is the unique stable fixed point of $\boldsymbol{f}$ in the unit cube $[0,1]^R$ different than $\mathbf{1}$ and the branching process is called *super-critical*. Therefore, with positive probability, the population will survive indefinitely.

(ii) If $m < 1$ then $\boldsymbol{\gamma} = \mathbf{1}$ is the unique stable fixed point of $\boldsymbol{f}$ in the unit cube $[0,1]^R$ and the branching process is called *sub-critical*. Therefore, with probability 1, the process will become extinct in finite time.

(iii) If $m = 1$ then $\boldsymbol{\gamma} = \mathbf{1}$ is the unique marginal fixed point of $\boldsymbol{f}$ in the unit cube $[0,1]^R$ and the branching process is called *critical*. Here, the expected time to extinction is infinite, despite the fact that extinction is bound to occur almost surely.

Unfortunately this theorem does not cover all the interesting cases, one important example for us being the phenotypic model for viral evolution. Nevertheless, one of the earliest results about decomposable branching processes is the generalization of the classification, due to Sevastyanov (see Harris [44] and Jiřina [48]). In the general decomposable case, there is a fourth alternative identified by Sevastyanov [44, 48] and in order to formulate this condition we need to introduce another important concept. A multitype Galton-Watson branching process is called *singular* if its probability generating function is linear without constant term: $\boldsymbol{f}(\boldsymbol{z}) = \boldsymbol{M}\boldsymbol{z}$. In this case, there is no branching since each particle produces exactly one particle that can be of any class and the process is equivalent to an ordinary finite Markov chain. More generally, a decomposable process may have *singular path components*. Two nodes i and j are said to be in same *path component* if there is a sequence of directed edges leading from i to j and a sequence of directed edges leading from j to i. This procedure defines a partition of the set of nodes into equivalence classes, called *path components* of the graph $\mathcal{G}(\boldsymbol{M})$. We say that a path component C of $\mathcal{G}(\boldsymbol{M})$ is a *singular path component*, if any particle whose class is in C has probability 1 of producing, in the next generation, exactly one particle whose class is in C. Equivalently, the component functions of the probability generating function corresponding to the classes in a path component C are linear functions of the variables corresponding to the classes in the path component C. In other words, the "part" of the probability generating function corresponding to the classes in C is that of a singular branching process. The existence of singular components is obviously an obstruction to extinction, for instance, in a decomposable singular process all path components are singular. In fact, the result of Sevastyanov states that if there is at least one *singular path component* then the branching process never become extinct, no matter what is the value of the malthusian parameter.

Example 2 *The graph corresponding to the general phenotypic model (FIG. 1 (b)) have two path components:* $\{0\}$ *and* $\{1, 2, 3, \ldots, R\}$. *In the simple phenotypic model (FIG. 1 (a)), the path components are exactly the sets containing one node,* $\{0\}$, $\{1\}$, $\ldots$, $\{R\}$. *From the expressions of the generating functions* (2) *and* (3) *it is clear that there are no singular path components in any of the models – simple or general, "with zero class" or "without zero class". Moreover, the general phenotypic model "without zero class" is* positively regular. *Therefore, the phenotypic model displays only the three regimes determined by the malthusian parameter, which depends on the values of the parameters* b, c, d *and* R.

It is important to stress that the regime of a multitype branching process can not be read from the mean matrix alone (i.e, the malthusian parameter). Essentially this happens because of the existence of decomposable branching processes with singular components.

Example 3 *Consider the following generating functions:*

$$\begin{aligned} \boldsymbol{g}(z,w) &= \left(1/2 + 1/2z^2, (dz+cw)^2\right) , \\ \boldsymbol{h}(z,w) &= \left(z, (dz+cw)^2\right) , \end{aligned}$$

where $0 < c, d < 1$ *and* $c + d = 1$. *They have the same mean matrix given by* $\boldsymbol{M} = \left(\begin{smallmatrix} 1 & 2d \\ 0 & 2c \end{smallmatrix}\right)$ *and so the malthusian parameter is* $m = \max\{1, 2c\}$. *It is easy to solve the fixed point equation* (8) *in both cases and compute the respective extinction probability vectors* (γ_1, γ_2): *for the function* $\boldsymbol{g}$ *we have that* $\gamma_1 = 1$ *and* $\gamma_2 = d^2/c^2$ *if* $0 \leqslant d \leqslant \frac{1}{2}$ *and* $\gamma_2 = 1$ *if* $\frac{1}{2} \leqslant d \leqslant 1$. *For the function* $\boldsymbol{h}$ *we have that* $\gamma_1 = \gamma_2 = 0$. *Therefore, the branching process defined by* $\boldsymbol{g}$ *becomes extinct if and only if* $c \leqslant 1/2$ *while the branching process defined by* $\boldsymbol{h}$ *never becomes extinct irrespective of the value of the malthusian parameter!*

C. Asymptotic Behaviour of Surviving Populations

According to the "Malthusian Law of Growth" it is expected that a super-critical branching process will grow indefinitely at a geometric rate proportional to m^n and we would like to write $\boldsymbol{Z}_n \approx m^n \boldsymbol{W}_n$, where $\boldsymbol{W}_n$ is a random vector with a finite "asymptotic distribution of classes" when $n \to \infty$. The formalization of this heuristic argument is due to Kesten and Stigum (see [49, 50] for the case of indecomposable multitype branching processes and [51] for the case of a general decomposable multitype branching processes).

Let us first recall the result in the indecomposable case (see Athreya and Ney [45]). Consider a super-critical branching process with $m > 1$ and suppose that the vector valued random variable $\boldsymbol{\zeta}$ satisfies a technical condition called *Kesten-Stigum "$\zeta \log \zeta$" condition* (see Lyons *et al.* [52] and Olofsson [53]), which is always satisfied in our case, since the probability distribution of the offsprings has finite support. It is natural to define the normalized random vector $\boldsymbol{W}_n = \boldsymbol{Z}_n/m^n$. This normalized random vector has a limit when $n \to \infty$, that is, there exists a scalar random variable $W \neq 0$ such that, with probability one,

$$\lim_{n\to\infty} \boldsymbol{W}_n = W\,\boldsymbol{u}\,,$$

where $\boldsymbol{u}$ is the normalized right eigenvector corresponding to the malthusian parameter m and

$$\mathbf{E}(W|\boldsymbol{Z}_0) = \boldsymbol{v}^{\mathrm{t}}\boldsymbol{Z}_0$$

where $\boldsymbol{v}$ is the left eigenvector corresponding to the malthusian parameter m.

An important step in the proof of Kesten-Stigum theorem is the Kurtz [54] *convergence of classes* theorem:

$$\lim_{n\to\infty} \frac{\boldsymbol{Z}_n}{|\boldsymbol{Z}_n|} = \boldsymbol{u} \qquad \text{(almost surely)}. \tag{9}$$

Combining this with the Perron-Frobenius theorem (see Meyer [43]) one obtains

$$\lim_{n\to\infty} \frac{\boldsymbol{Z}_n}{|\boldsymbol{Z}_n|} = \lim_{n\to\infty} \frac{\langle \boldsymbol{Z}_n\rangle}{|\langle \boldsymbol{Z}_n\rangle|} = \boldsymbol{u}\,, \tag{10}$$

where the convergence in the first limit is in probability. The approach adopted in [29–31] relates only to the second equality involving the limit of mean values of equation (10). By explicitly considering the microscopic model as a multivariate branching process the equality of the two limits in equation (10) is guaranteed. This result may be useful for computational simulations of the model, since one may compute the eigenvector $\boldsymbol{u}$ by sampling the population and taking averages.

The meaning of the Kesten-Stigum theorem is that the total size of the population divided by m^n, converges to a random vector, but the relative proportions of the various "classes" approach fixed limits. Since we are assuming that the process is indecomposable the normalized right eigenvector $\boldsymbol{u} = (u_0, \ldots, u_R)$ is positive and is normalized so that $\sum_r u_r = 1$, therefore it defines a probability distribution on the set of classes $\{0, \ldots, R\}$. It is called the *asymptotic distribution of classes* of the multitype branching process.

In order to extend these results to the case where the branching process is decomposable one should employ the *Frobenius normal form* of the mean matrix $\boldsymbol{M}$, which is reducible in this case (see Gantmatcher [42]). Kesten and Stigum [51] shows that it is possible, by rearranging the rows and columns, to rewrite the mean matrix in a block upper triangular form in such a way that the diagonal blocks are irreducible square matrices associated to components of the decomposable branching process. By a *component* of a decomposable branching process we mean a subset of classes such that their associated nodes in the graph $\mathcal{G}(\boldsymbol{M})$ forms a path component. Let $\{C_k \ : \ 0 \leqslant k \leqslant N\}$ be the set of components of $\mathcal{G}(\boldsymbol{M})$ ordered according to which $C_k \prec C_l$ if there is a sequence of directed edges leading from some $i \in C_k$ to some $j \in C_l$. Given two components C_k and C_l define the sub-matrix

$$\boldsymbol{M}(k,l) = M_{ij} \quad \text{with} \quad i \in C_k\,, j \in C_l\,.$$

Then, for each k, the square sub-matrix $\boldsymbol{M}(k) = \boldsymbol{M}(k,k)$ is the irreducible mean matrix of the sub-process

$$\boldsymbol{Z}_n(k) = \{Z_n^i \ : \ i \in C_k\}\,.$$

Now the order of the components C_k allows us to rearrange the rows and columns of $\boldsymbol{M}$ in such a way that

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{M}(0) & \boldsymbol{M}(0,1) & \boldsymbol{M}(0,2) & \ldots & \boldsymbol{M}(0,N) \\ 0 & \boldsymbol{M}(1) & \boldsymbol{M}(1,2) & \ldots & \boldsymbol{M}(1,N) \\ 0 & 0 & \boldsymbol{M}(2) & \ldots & \boldsymbol{M}(2,N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \boldsymbol{M}(N) \end{pmatrix} \tag{11}$$

Therefore, the sub-process $\boldsymbol{Z}_n(k)$ "receives input" from the sub-process $\boldsymbol{Z}_n(l)$, with $k < l$, throughout the sub-matrix $\boldsymbol{M}(k,l)$. Note that if the sub-matrices $\boldsymbol{M}(k,l)$ are all zero then the branching process splits as a sum of independent indecomposable branching processes.

Example 4 *The matrices* (5) *and* (6) *of the simple and the general phenotypic models, respectively, already are in the normal form* (11)*:*

(i) In the simple phenotypic model we have that $N = R$, with one-dimensional diagonal sub-matrices $\boldsymbol{M}(k) = \big(k(1-d)\big)$, with one-dimensional sub-matrices $\boldsymbol{M}(k,k+1) = \big((k+1)d\big)$ and one-dimensional sub-matrices $\boldsymbol{M}(k,l) = 0$ if $l > k+1$.

(ii) In the general phenotypic model we have $N = 1$, with the first diagonal sub-matrix $\boldsymbol{M}(0) = \big(0\big)$ (or $\boldsymbol{M}(0) = \big(1\big)$ for the first variation of the model), the second diagonal sub-matrix $\boldsymbol{M}(1) = M_{ij}$, with $i, j = 1, \ldots, R$ and $\boldsymbol{M}(0,1) = \big(d\ 0 \cdots 0\big)$.

Now observe that if $Z_0^i = 1$ with $i \in C_k$ then for $l > k$, the sub-process $\boldsymbol{Z}_n(l) = \boldsymbol{0}$ for all $n \geqslant 0$. That is, the branching process behaves as if the sub-processes $\boldsymbol{Z}_n(l)$ for all $l > k$ did not exist. Since each non-zero diagonal sub-matrix $\boldsymbol{M}(l)$ is irreducible, it has a largest positive eigenvalue $m(l)$ and then we may define the *effective malthusian parameter* of the sequence of sub-processes $(\boldsymbol{Z}_n(0), \ldots, \boldsymbol{Z}_n(k))$ to be

$$m_{\mathrm{e}}(k) = \max_{l \leqslant k}\{m(l)\}\,.$$

The simplest case is when all $m(l)$ are simple eigenvalues of their respective sub-matrices $\boldsymbol{M}(l)$ – they are distinct amongst each other – this is exactly the case for matrices (5) and (6).

In Kesten and Stigum [51] the result about the asymptotic behaviour of irreducible super-critical branching process is generalized to the reducible case. The main theorem applied to the case where all $m(l)$ are simple eigenvalues of their respective sub-matrices $\boldsymbol{M}(l)$ states that if the effective malthusian $m_{\mathrm{e}}(k) > 1$ and the "$\zeta \log \zeta$" condition holds then for the normalized random vector $\boldsymbol{W}_n(k) = \boldsymbol{Z}_n/(m_{\mathrm{e}}(k))^n$ there exists a scalar random variable $W \neq 0$ such that, with probability one,

$$\lim_{n\to\infty} \boldsymbol{W}_n(k) = W\,\boldsymbol{u}(k)\,,$$

where $\boldsymbol{u}(k)$ is the normalized right eigenvector corresponding to the effective malthusian parameter $m_{\mathrm{e}}(k)$ and

$$\mathbf{E}(W|\boldsymbol{Z}_0) = \boldsymbol{v}(k)^{\mathrm{t}}\boldsymbol{Z}_0\,,$$

where $\boldsymbol{v}(k)$ is the left eigenvector corresponding to the effective malthusian parameter $m_{\mathrm{e}}(k)$. Moreover, Kurtz's *convergence of classes* theorem (9) still holds. But one should note that the normalized right and left eigenvectors are not positive anymore. In fact, $\boldsymbol{v}(k)$ may have negative entries, but only those associated to the components C_l with $l \leqslant k$, for which $\boldsymbol{Z}_0$ is zero. The right normalized eigenvector is of the form $\boldsymbol{u}(k) = (u_0, \ldots, u_r, 0, \ldots, 0)$, where $(u_0, \ldots, u_r)$ is the non-negative right normalized eigenvector of the sub-matrix corresponding to the the sequence of sub-processes $(\boldsymbol{Z}_n(0), \ldots, \boldsymbol{Z}_n(k))$, and so is a probability distribution.

D. Critical Behavior and Regime Transition

The critical state separates the super-critical and the sub-critical regimes where the branching process has two distinct behaviors in time and thus characterizes the existence of regime transition with genuine critical behavior. In fact, the decay of correlation functions described in the next section for the case of the simplest model clarifies this point.

Although in a critical branching process $\boldsymbol{Z}_n \to 0$, almost surely, when $n \to \infty$, one still may obtain a meaningful asymptotic law by conditioning on non-extinction. See Mullikin [55] and Joffe and Spitzer [56] for the indecomposable case and Foster and Ney [57] for certain decomposable cases.

In the indecomposable critical case $\boldsymbol{Z}_n$ grows at a linear rate proportional to n (see Harris [44] or Athreya and Ney [45]), and so one should consider the normalized random vector $\boldsymbol{Y}_n = \boldsymbol{Z}_n/n$. If the second moments are finite and the branching process is non-singular, there is a scalar random variable $Y \neq 0$ such that

$$\lim_{n\to\infty} \boldsymbol{Y}_n = Y\boldsymbol{u} \quad \text{given that } \boldsymbol{Z}_n \neq 0\,,$$

where $\boldsymbol{u}$ is the normalized right eigenvector corresponding to the malthusian parameter m and with convergence only in *distribution*, which is weaker than the *almost surely convergence* in the super-critical case.

IV. THE SIMPLE PHENOTYPIC MODEL

For the simple phenotypic model it is easy to compute the eigenvalues λ_r of the mean matrix $\boldsymbol{M}$:

$$\lambda_r = rc = r(1-d) \qquad r = 0, \ldots, R\,.$$

In particular, the malthusian parameter is the largest positive eigenvalue

$$m = \varrho(\boldsymbol{M}) = \lambda_R = Rc = R(1-d)\,. \tag{12}$$

Therefore we have the following immediate result.

Theorem 1 *The simple phenotypic model has three distinct regimes.*

(i) If $R(1-d) < 1$ then the branching process is sub-critical. *That is, with probability* 1*, the virus population becomes extinct in finite time.*

(ii) If $R(1-d) > 1$ then the branching process is super-critical. *That is, with positive probability, the virus population survives and grows indefinitely at an exponential rate proportional to m^n when $n \to \infty$.*

(iii) If $R(1-d) = 1$ then the branching process is critical. *That is, with probability* 1*, the virus population becomes extinct but this may take an infinite time to happen.*

Proof. This is a straightforward consequence of the main result about the classification of a multitype branching process and equation (12). □

Thus theorem 1 provides a partition of the parameter space $\{(d, R) : d \in [0, 1],\, R \in \mathbb{N}\}$ of the simple phenotypic model into two regions (see FIG. 2).

It is also important, specially in order to describe the asymptotic behaviour in the super-critical case, to know the left eigenvectors $\boldsymbol{v}$ and right eigenvectors $\boldsymbol{u}$ corresponding to the eigenvalue λ_R Let us write the left and right eigenvectors in components as

$$\boldsymbol{v} = (v_0, v_1, \ldots, v_R) \quad \text{and} \quad \boldsymbol{u} = (u_0, u_1, \ldots, u_R)$$

and satisfying the normalization conditions (7).

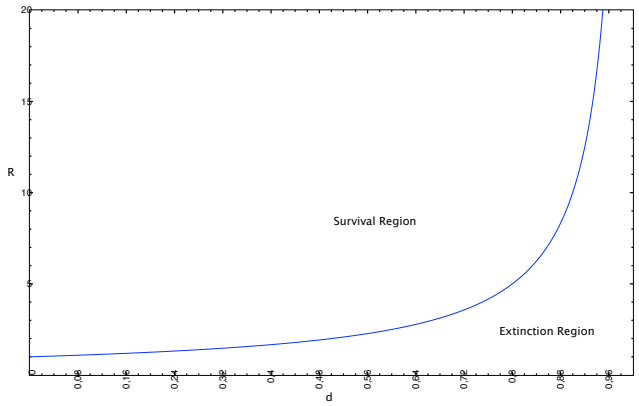

FIG. 2. Graph of the function $R = 1/(1-d)$ (in blue). The region below this curve corresponds to the sub-critical parameters (d, R) and the region above this curve corresponds to the super-critical parameters (d, R). The curve itself corresponds to the critical parameters (d, R).

(i) In the version "with zero class" the left eigenvector $\boldsymbol{v}$ is given by

$$\boldsymbol{v} = \frac{1}{(1-d)^R}\,(0,\ldots,0,1)$$

and the right eigenvector $\boldsymbol{u}$ is given by

$$u_k = \binom{R}{k}(1-d)^k\, d^{R-k} = \mathrm{binom}(k; R, 1-d)\,,$$

with $k = 0, 1, \ldots, R$.

(ii) In the version "without zero class" there is no components v_0 and u_0. The left eigenvector $\boldsymbol{v}$ is given by

$$\boldsymbol{v} = \frac{1-d^R}{(1-d)^R}\,(0,\ldots,0,1)$$

and the right eigenvector $\boldsymbol{u}$ is given by

$$u_k = \frac{1}{1-d^R}\binom{R}{k}(1-d)^k\, d^{R-k}\,,$$

with $k = 1, \ldots, R$.

It is interesting to note that the simple phenotypic model is a "completely solvable" branching process in the sense that we may explicitly solve the spectral problem for its mean matrix independently of the numerical values of the parameters.

Next we turn to the computation of the extinction probabilities γ_r. In this case, it is necessary to solve a non-linear system of polynomial equations:

$$\begin{aligned} z_0 &= 1 \\ z_1 &= dz_0 + (1-d)z_1 \\ z_2 &= \big(dz_1 + (1-d)z_2\big)^2 \\ &\vdots \\ z_R &= \big(dz_{R-1} + (1-d)z_R\big)^R \end{aligned} \tag{13}$$

This may be done in a recursive way, since the equation for z_0 is already solved $z_0 = 1$ and the equation for z_k depends only on z_k and z_{k-1}. Thus we get for $R = 0, 1, 2$:

$$\gamma_0 = 1$$
$$\gamma_1 = 1$$
$$\gamma_2 = \begin{cases} d^2/(1-d)^2 & \text{for} \quad 0 \leqslant d \leqslant \frac{1}{2} \\ 1 & \text{for} \quad \frac{1}{2} \leqslant d \leqslant 1 \end{cases}$$

When $R \geqslant 3$ the formulas become very complicated and when $R \geqslant 5$ the equation may not even be solvable by radicals, but in general one may write

$$\gamma_r = \begin{cases} f(d) & \text{for} \quad 0 \leqslant d \leqslant d_c \\ 1 & \text{for} \quad d_c \leqslant d \leqslant 1 \end{cases}$$

where $d_c = \frac{r-1}{r}$ and $f(d)$ is a strictly increasing smooth function on $[0, 1[$ satisfying: (i) $f(0) = 0$, (ii) $f(d_c) = 1$, (iii) $f(d) < 1$ for $0 \leqslant d < d_c$ and (iv) $\lim_{d\to 1} f(d) = +\infty$. This expression suggests that the surviving probabilities $\omega_r = 1 - \gamma_r$ can be interpreted as an *order parameter* associated to the occurrence of a *phase transition* when the deleterious probability d attains the critical point $d_c = \frac{r-1}{r}$, which marks the transition from super-criticality to sub-criticality,

$$\omega_r = \begin{cases} g(d) & \text{for} \quad 0 \leqslant d \leqslant d_c \\ 0 & \text{for} \quad d_c \leqslant d \leqslant 1 \end{cases}$$

where $g(d) = 1 - f(d)$ and thus satisfies: (i) $g(0) = 1$, (ii) $g(d_c) = 0$, (iii) $g(d) > 0$ for $0 \leqslant d < d_c$ and (iv) $\lim_{d\to 1} g(d) = -\infty$. Observe that for a fixed numerical value of d, the system of equations (13) can be easily solved by numerical approximation using Newton's method. For instance, in FIG. 3 we show the curves for the surviving probabilities ω_r as functions of d for $r = 2, 3, 4, 5, 6, 7$.

The result shows that, with respect to ω_r, the model has a critical behavior in complete analogy to a second order phase transition (see FIG. 3). Therefore, the critical properties of the model can be characterized by means of relevant critical exponents.

Finally, it is not difficult to see that for fixed d, the numbers γ_r satisfy $1 \geqslant \gamma_2 \geqslant \gamma_3 \geqslant \ldots \geqslant \gamma_R$ and therefore the extinction probability for a general initial condition $\boldsymbol{Z}_0 = (Z_0^0, \ldots, Z_0^R)$ may be estimated far from the critical deleterious probability $d_c = (R-1)/R$ by

$$\mathbf{P}(\boldsymbol{Z}_n = 0 \text{ for some } n | \boldsymbol{Z}_0) \;\approx\; \gamma_2^{|\boldsymbol{Z}_0'|}\,, \tag{14}$$

where $|\boldsymbol{Z}_0'| = Z_0^2 + \ldots + Z_0^R$ and near $d_c = (R-1)/R$ by

$$\mathbf{P}(\boldsymbol{Z}_n = 0 \text{ for some } n | \boldsymbol{Z}_0) \;\approx\; \gamma_R^{Z_0^R}\,. \tag{15}$$

It has been demonstrated that large population passages are able to increase the adaptability of virus populations [10]. On the other hand, small population passages represented by bottleneck events are capable to

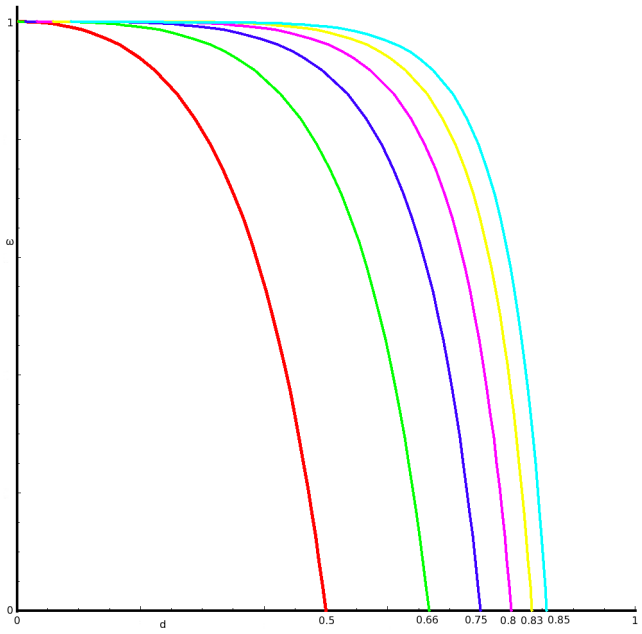

FIG. 3. Curves for the surviving probability $\omega_r(d)$ as function of d for $r = 2, 3, 4, 5, 6, 7$.

increase the risk towards viral extinction. Among the aspects of abrupt population reductions are the exacerbated effects of drift that coupled with the Muller's hatchet principle [58] may lead to the random and progressive lost of the best adapted virus in a population. It also has been suggested that large virus populations bearing a significant phenotypic diversity are more adaptable to environment fluctuations and robust. It is correct to assume that large initial virus populations colonizing new hosts may show better survival probabilities than populations recovering from bottlenecks. In this way the size of the viral innoculums may have an impact in the survival rates of different virus populations.

From now on we shall split the analysis of the simple phenotypic model according to which it is sub-critical, super-critical or critical.

It is important to note that the existence of a clear cut between regimes of survival and non survival populations by means of a critical state is directly related to the problem of lethal mutagenesis for viral populations. In fact, proposition (i) in Theorem 1 is precisely the Bull, Sanjuán and Wilke conjecture [12] also represented in FIG. 2.

A. The Sub-critical Regime: Lethal Mutagenesis

The first consequence of theorem 1 is a proof, in the context of this model, of the conjecture of Bull, Sanjuán and Wilke for *lethal mutagenesis* [12].

Corollary 2 *In the simple phenotypic model, the virus population becomes extinct in finite time, with probability* 1*, if the product of the* neutral effect probability $(1-d)$ *with the* maximum replication capacity R *is strictly less than one:*

$$(1-d)R < 1\,.$$

The main conclusion here is that the existence of lethal mutagenesis depends on "genetic components" (mutational rates) and other additional deleterious effects (host driven pressures intensifications), as well as on strict "ecological components", namely, the maximum replication capacity of the particles in the population and on the initial population size. As a result the viral population may reach extinction by increasing the number of deleterious mutations per replication cycle, by decreasing the value of R in the population or by a combination of the two mechanisms. The mutational strategy is the basis of treatments using mutagenic drugs (see Crotty *et al.* [8]) that induce errors in the generation process of new viral particles reducing their replication capacity. A straightforward consequence of corollary 2 is that a single particle showing the maximum replication capacity R is able to rescue a viral population driven to extinction by mutagenic drugs. If it is assumed that RNA virus populations correspond to a swarm of variants with distinct replication capacities, for a therapy to become effective it is important that it will eliminate the classes represented by particles with highest replication capacities. As a conclusion the higher the replication capacity of the first particles infecting the organism the larger should be the number of deleterious mutations (or effects) and therefore the larger should be the drug concentration. This can be a clear limitation for treatments based on mutagenic drugs.

B. The Super-critical Case: Relaxation and Equilibrium

In the super-critical regime, the population grows at a geometric pace indefinitely. Nevertheless, there are two distinct phases that occur during this growth: a transient phase ("relaxation"or "recovery time") and a dynamical stationary phase.

1. *Relaxation towards equilibrium.*

An important question concerning the adaptation process of a viral population to the host environment is the typical time needed to achieve the equilibrium state. As the equilibrium is characterized by constant mean replication capacity an obvious criteria to measure the time to achieve equilibrium would be by the vanishing variation of this variable as used in other studies (Aguirre *et al.* [14]). Nevertheless, this method is clearly subjected

to the limitations of numerical accuracy with evident disadvantages if one wants a sharp and universal criterion to differentiate populations from the point of view of how fast a population can be typically stabilized in a organism.

Viral populations are commonly submitted to transient regimes. As pointed out earlier the infection transmission process represents the passage of a small number of particles from one organism to another in such a way that in this process the viral population is submitted to a subsequent *bottle-neck effect* during spreading of viruses in the host population. In order to approach the problem of relaxation after a bottleneck process in a more sound basis the natural quantity to be considered is the characteristic time derived from the decay of the mean auto-correlation function. The temporal correlation function $C(n)$ is typically of the form $\exp(-\alpha n)$ and the decay rate is given by the parameter α. The natural way to define a characteristic time T to achieve equilibrium is by setting $T = 1/\alpha$. In order to find the characteristic decay rates one should consider the recursive application of the mean matrix $\boldsymbol{M}$ on the initial population: $\boldsymbol{Z}_0^{\mathrm{t}}\,\boldsymbol{M}^n\,\boldsymbol{Z}_0$. In fact, it is enough to consider the canonical initial population $\boldsymbol{Z}_0 = \boldsymbol{e}_R = (0, 0, ..., 1)$. By direct inspection it is easily verified that the decay of correlations is typically exponential and given by

$$C(n) = \exp\big(-\log(R(1-d))\,n\big)\,,$$

where $m = R(1-d)$ is the malthusian parameter. The decay rate is therefore given by $\alpha = \log(R(1-d))$.

Among others, one possible application of this result relates to the very initial phase of the infection process. If we consider that during this phase the host immune system has not been yet stimulated against the virus, one can assume that the deleterious effects would be solely represented by the viral intrinsic mutation rates. Therefore, the largest the value of R, i.e., the largest the replication capacity of the initial viral particle the fastest the progeny auto-correlation decays and reaches equilibrium stabilizing the viral population; intuitively the parameter R defines the degree of virulence of the infection during the early stage of the infective process. The increment of deleterious effects plays an opposite role on the decay rates. In fact, as it will be shown below the closest the parameter d is to its critical value d_c more time is needed to achieve equilibrium.

2. The Dynamical Stationary State.

When the simple phenotypic model is super-critical and is initialized with exactly one particle in the class r ($Z_0^r = 1$) the effective malthusian parameter is $m_{\mathrm{e}} = \lambda_r = rc = r(1-d)$ with corresponding normalized right eigenvector $\boldsymbol{u}(r) = \big(u_0(r), \ldots, u_r(r), 0, \ldots, 0\big)$, where the components $u_k(r)$, with $k = 0, \ldots, r$, are

$$u_k(r) = \mathrm{binom}(k; r, 1-d) = \binom{r}{k}(1-d)^k\,d^{r-k}\,. \quad (16)$$

Therefore, the simple phenotypic model has $R-1$ distinct asymptotic distributions of types of particles, describing $R-1$ distinct *dynamical stationary states*, characterized by their *asymptotic distribution of classes* given by (16) (up to a random scalar perturbation), each one of these being achieved when the branching process is initialized with exactly one particle in the class r ($Z_0^r = 1$) for $r = 2, \ldots, R$, respectively. Note that when $r = 0, 1$ the process is always sub-critical.

Theorem 3 *If the simple phenotypic model is super-critical with malthusian parameter $m = R(1-d)$ and starts with at least one particle of class R then, in the long run, the relative number of particles in each class reaches a stable stationary dynamical state and is (up to a random scalar perturbation) distributed according to the* Binomial Distribution: $\mathrm{binom}(k; R, 1-d)$*, where $k = 0, \ldots, R$ are the replication classes.*

Proof. This is consequence of Kesten-Stigum results about the asymptotic behaviour of super-critical multitype branching processes and the computation of the normalized right eigenvector associated to the malthusian parameter $m = R(1-d)$ given by equation (16). □

From theorem 3 we immediately obtain:

- The *mean replication capacity* is

$$\mathbf{E}(\boldsymbol{u}) = R(1-d)\,.$$

- The *phenotypic diversity* is

$$\mathbf{Var}(\boldsymbol{u}) = Rd(1-d)\,.$$

It is well accepted that the phenotypic diversity is an important characteristic of the viral population intuitively related to the idea of population robustness [59, 60]. In fact, a homogeneous population would be less flexible from the point of view of adaptation. The variance associated with the stationary state can be understood as a natural quantity to measure diversity. It shows that the maximum value of the phenotypic diversity $r/4$ is reached if $d = 1/2$ for any value of r. If $R > 2$ the variation of the phenotypic diversity as a function of d shows that there are two different domains to be considered: below $d = 1/2$ the diversity is an increasing function of d. It implies that if the population has a typical value of $d < 1/2$ the effect of inducing an increment of d (for instance using mutagenic drugs) increases the phenotypic diversity. For $1/2 < d < d_c$ this effect reverses and diversity decreases with increasing d. This result raises the question if in normal conditions the viral population adapt to the host environment guided by a principle of maximum phenotypic diversity or if the environmental conditions simply contribute to fix one possible value of diversity for the population that may vary from one to another host organism. Interesting enough, the natural deleterious mutations has been measured for certain viruses and, as shown in the TABLE I, they are close to

the value $d = 1/2$. In the first case one could preview that the set point of the viral disease should be invariant (or with small variation) for all hosts. On the other hand the second hypothesis leads to the idea of different responses to treatment depending on the initial value of d before the adoption of treatment strategies to improve d. At the present the two scenarios may apply to different type of viruses and this point clearly has to be decided experimentally.

Virus	U_d	$(1-d) = \mathrm{e}^{-U_d}$	REF.
VSV	0.692	0.500	[21]
TEV	0.773	0.461	[37]
ΦX174	0.72 - 0.77	0.48 - 0.46	[61]
Qβ	0.74 - 0.86	0.47 - 0.42	[61]

TABLE I. Experimental results of deleterious mutation rates: (VSV) vesicular stomatitis virus, (TEV) Tobacco etch virus and (ΦX174, Qβ) bacterial viruses.

Another important consequence of the above results concerns the efficiency of the use of mutagenic drugs. In the region $d < 1/(R+1) < 1/2$ the viral population's most representative particle is the fittest one (class R). If we assume that the drug action is deeply influenced by drug transport coefficients in different host tissues, it is important to be assured that local drug concentrations will still eliminate the set of class R particles. If d increases beyond $1/(R+1)$ the representative particle of the population is not anymore the fittest one but a set of particles from different replication classes. Therefore the main drug target represents a group of average replicating particles of a population with higher phenotypic diversity in which resistance drug mutants can be contained. In this case one would say that the viral population displays a kind of endogenous strategy to scape the deleterious action of the mutagenic drug. If we assume that deleterious effects are small in the early stage of the infection process we should expect that at this stage the drug efficiency would be maximum reinforcing the successful practice of post exposure therapy, currently adopted in the case of HIV infections [62].

C. The Critical Case: Extinction Threshold

The clearest way to characterize the time behavior of the viral population at or around the critical point is through the typical time T to approach equilibrium derived from the decay of correlations described above. The expression $T = 1/\log(R(1-d))$ shows that at the critical point the equilibrium state is never reached, i.e., the decay to equilibrium is at least non-exponential. A scaling exponent characterizing the behavior of T in the neighborhood of the critical point d_c can be easily obtained. The expansion around $d_c = (R-1)/R$ gives

$$T \approx (1-d_c)\,|d-d_c|^{-1}\,.$$

Although it is always possible to calculate intermediate distributions of progeny, it is quite easy to see that at the critical point the time evolution of densities never achieves an invariant density. Unlike in the super-critical regime, the relative number of particles in each class/subpopulation is never stable. Nevertheless, our preliminary results concerning the dynamics of fluctuations show that the time variation of the numbers of particles in each separated class follows a pattern such that the variation observed in one class is rigorously the same observed in all the others. This indicate a high level of correlation between the classes in complete analogy with critical phenomena of many physical systems. We conjecture that in the critical regime the highly correlated classes in the population behave as an inseparable whole such that the notion of the population divided in separated classes becomes meaningless. In other words the correlation between classes makes them behave as if they constitute one unique class, which reminds one of the basic properties of the error threshold in Eigen's theory [63]. In fact, according to Eigen, when mutational rates are increased beyond a threshold, infinite viral populations are not anymore able to retain its best adapted variants. At this critical mutation level, selection is overruled by mutation and all variants share the same fitness status. Moreover, populations at Eigen's error threshold do not become extinct, but well defined replication classes cease to exit, as particles hazardously wander through the surface of a flat landscape. If in the super-critical case the notion of the mean replication capacity and therefore that of the "mean viral particle" exists defining a typical scale in the system, in the critical case this notion is absent. Therefore, in using branching processes to model the time behavior of viral populations the concept of error threshold should be identified with that of criticality. In the same direction of reasoning, in terms of branching processes the existence of lethal mutagenesis should be identified with that of critical behavior of the model. It is worth to note that there is a correspondence between Eigen's model of molecular evolution and the equilibrium statistical mechanics of an inhomogeneous Ising system (see Leuthäusser [64]). For further discussions about the error threshold in connection to extinction see (see [65, 66]).

The critical behavior of the model can also be observed through the survival probability ω for $d \lesssim d_c$ as show in FIG. 3. Expansion of the survival probability around the critical point by means of functional equation (eq. (13)) gives directly

$$\omega \approx 2\frac{R}{d_c}\,|d-d_c|\,.$$

It is interesting to note that the critical exponents of T and ω are the same found in critical behavior of a large class of dynamical random networks (see Riordan and

Warnke [67]. This fact is reminiscent from the deep relation existing between branching process and random network theory, where the survival probability function of a branching process is identified with the order parameter associated to the emergence of the giant cluster in a dynamical random network. This fundamental observation goes back to Karp [68] and more recently it has become the central technique in the study of more general models of random networks [69]. The relation between the two theories is certainly expected to bring important results in the future.

Finally, it is noteworthy that here we talk about criticality of a process taking place in time, and therefore the term *critical phenomenon* (imported from equilibrium statistical mechanics of space distributed systems) is used to highlight the fact that the survival probability behaves like an order parameter and the amount of deleterious effects quantified by d behaves as a control parameter that can be changed by external means.

V. CONCLUSIONS AND OUTLOOK

Using the previous theoretical model for virus evolution proposed by Lázaro *et al.* [10] and Aguirre *et al.* [13, 14] as a starting point we show that virus evolution can be described by an exact solvable multivariate branching process. By applying our approach we are able to identify crucial aspects of the dynamics of replicating viral populations on a sound theoretical basis. Among these several aspects we are able to demonstrate that – as long as the beneficial effects are close to zero – the two main driving features of a virus population are the maximum replication capacity and the fraction of the population not affected by deleterious effects. Based on this result we show that, as proposed by Bull *et al.*, if the product between the above mentioned parameters $m = R(1-d)$ yields a value less than one the population undergoes extinction. On the other hand, if $m = R(1-d)$ is greater than one and the environment is constant we show that the population will reach an asymptotic stationary state characterized by the stability of the replicative classes. However, the time to reach the stationary equilibrium strictly depends on how intense is the deleterious effect, more precisely, the higher d the longer is the transient phase and when d approaches its critical value d_c the transient tends to infinity.

According to our explicit formulas for the progeny distribution, we demonstrate that virus populations maximize their phenotypic diversity by replicating with d near $1/2$, for any value of R. We speculate that this might be a universal property for RNA viruses that replicate under high mutational rates. In this way by increasing their phenotypic diversity viruses augments their chances of survival escaping and adapting to environmental pressures. Maintenance of high mutation rates makes it difficult for a population to retain their best replicative classes. As a consequence, the best adapted classes are not the most represented ones in the population, thus not characterizing a classical Darwinian evolution process. As far as branching process modeling of viral evolution is concerned its critical behavior partially resembles the concept of error threshold in Eigen's theory of molecular quasispecies. In this regime the replicative classes lose their independence in the sense that they become so much correlated that the whole set of classes behaves as a single one.

We also demonstrate that by keeping the deleterious effects constant the survival probability of a virus population will depend on its initial population size. By increasing the population size at time zero we push the survival probability curves, in the region before the critical point, towards one (see FIG. 3 and equations (14) and (15)). According to this result it can be speculated that virus with greater innoculums have a better chance of survival colonizing new hosts. Interestingly enough and in a frontal disagreement to the above observations it has been shown that only a limited number of particles, and in some cases even one particle, is enough to start a new infectious process in a host [19, 20]. However, according to the model and as discussed before, the R parameter determines the success of an incoming virus population because the corresponding value of d_c is uniquely given by R. The present work suggests that minimum innoculums must have at least one particle with replicative capacity large enough in order to survive in the new host. We speculate that those particles with maximum replicative capacity should constitute the effective innoculum described in Zwart *et al.* [19]. In fact, the experimental data about viral load in HIV early infected patients strongly suggests that the host deleterious effects over the viral population are minimal and increase after the onset of the immunological response [70]. We note that the characteristic form of this data can be easily reproduced by the model (see Castro [71] and manuscript in preparation [72]).

Finally, it is important to mention that the close relation of the theory of branching processes (as used in the present work) and dynamical Erdös-Renyi type networks indicates that the latter may be brought to bear in the modeling of virus populations. The relation between these two theories is undoubtedly a research avenue with promising potential to improve our knowledge of the dynamical laws governing the evolution and adaptation of viral populations.

ACKNOWLEDGMENTS

FA wish to acknowledge the support of CNPq through the grant PQ-313224/2009-9. FB recieved support from the Brazilian agency FAPESP. DC received financial support from the Brazilian agency CAPES.

[1] E. Domingo and *et al.*, "The quasispecies (extremely heterogeneous) nature of viral rna genome populations: biological relevance – a review," Gene **40**, 1–8 (1985).

[2] E. Domingo, E. Baranowski, C. M. Ruiz-Jarabo, A. M. Martín-Hernández, J. C. Sáiz, and C. Escarmís, "Quasispecies structure and persistence of rna viruses," Emerg. Infect. Dis. **4**, 521–527 (1998).

[3] D. A. Steinhauer, E. Domingo, and J. J. Holland, "Lack of evidence for proofreading mechanisms associated with an rna virus polymerase," Gene **122**, 281–288 (1992).

[4] S. F. Elena and R. Sanjuán, "Adaptive value of high mutation rates of rna viruses: separating causes from consequences," J. Virology **79**, 11555–11558 (2005).

[5] C. Escarmís, E. Lázaro, and S. C. Manrubia, "Population bottlenecks in quasispecies dynamics," Curr. Top. Microbiol. Immunol. **299**, 141–170 (2006).

[6] S. Ojosnegros, N. Beerenwinkel, and E. Domingo, "Competition-colonization dynamics: An ecology approach to quasispecies dynamics and virulence evolution in rna viruses," Commun. Integr. Biol. **3**, 333–336 (2010).

[7] S. Ojosnegros, N. Beerenwinkel, T. Antal, M. A. Nowak, C. Escarmís, and E. Domingo, "Competition-colonization dynamics in an rna virus," Proc. Natl. Acad. Sci. U. S. A. **107**, 2108–2112 (2010).

[8] S. Crotty, C. E. Cameron, and R. Andino, "RNA virus error catastrophe: direct molecular test by using ribavirin," Proc. Natl. Acad. Sci. U. S. A. **98**, 6895–6900 (2001).

[9] C. O. Wilke, J. L. Wang, C. Ofria, R.E Lenski, and C. Adami, "Evolution of digital organisms at high mutation rates leads to survival of the flattest," Nature **412**, 331–333 (2001).

[10] E. Lázaro, C. Escarmís, E. Domingo, and S. C. Manrubia, "Modeling viral genome fitness evolution associated with serial bottleneck events: Evidence of stationary states of fitness," J. Virology **76**, 8675–8681 (2002).

[11] S. C. Manrubia, E. Lázaro, E. Pérez-Mercader, C. Escarmís, and E. Domingo, "Fitness distributions in exponentially growing asexual populations," Phys. Rev. Lett. **90**, 188102 (2003).

[12] J. J. Bull, R. Sanjuán, and C. O. Wilke, "Theory of lethal mutagenesis for viruses," J. Virology **18**, 2930–2939 (2007).

[13] J. Aguirre and S. C. Manrubia, "Effects of spatial competition on the diversity of a quasispecies," Phys. Rev. Lett. **100**, 038106 (2008).

[14] J. Aguirre, E. Lázaro, and S. C. Manrubia, "A trade-off between neutrality and adaptability limits the optimization of viral quasispecies," J. Theor. Biol. **261**, 148–155 (2009).

[15] J. W. Drake and J. J. Holland, "Mutation rates among RNA viruses," Proc. Natl. Acad. Sci. U. S. A. **96**, 13910–13913 (1999).

[16] R. Sanjuán, M. R. Nebot, N. Chirico, L. M. Mansky, and R. Belshaw, "Viral mutation rates," J. Virology **84**, 9733–9748 (2010).

[17] E. Batschelet, E. Domingo, and C. Weissmann, "The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate," Gene **1**, 27–32 (1976).

[18] J. M. Cuevas, F. González-Candelas, A. Moya, and R. Sanjuán, "The effect of ribavirin on the mutation rate and spectrum of Hepatitis C virus in vivo," J. Virology **83**, 5760–5764 (2009).

[19] M. P. Zwart, J.-A. Daròs, and S. F. Elena, "One is enough: In vivo effective population size is dose-dependent for a plant RNA virus," PLoS Pathog **7**, e1002122 (2011).

[20] B.F. Keele and *et al.*, "Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection," Proc. Natl. Acad. Sci. U. S. A. **105**, 7552–7557 (2008).

[21] R. Sanjuán, A. Moya, and S. F. Elena, "The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus," Proc. Natl. Acad. Sci. U. S. A. **101**, 8396–8401 (2004).

[22] J. J. Holland, E. Domingo, J. C. de la Torre, and D. A. Steinhauer, "Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis," J. Virology **64**, 3960–3962 (1990).

[23] C. H. Lee, D. L. Gilbertson, I. S. Novella, R. Huerta, E. Domingo, and J. J. Holland, "Negative effects of chemical mutagenesis on the adaptive behavior of vesicular stomatitis virus," J. Virology **71**, 3636–3640 (1997).

[24] L. A. Loeb, J. M. Essigmann, F. Kazazi, J. Zhang, K. D. Rose, and J. I. Mullins, "Lethal mutagenesis of HIV with mutagenic nucleoside analogs," Proc. Natl. Acad. Sci. U. S. A. **96**, 1492–1497 (1999).

[25] S. Sierra, M. Dávila, P. R. Lowenstein, and E. Domingo, "Response of foot-and-mouth disease virus to increased mutagenesis: influence of viral load and fitness in loss of infectivity," J. Virology **74**, 8316–8323 (2000).

[26] A. Grande-Pérez, S. Sierra, M. G. Castro, E. Domingo, and P. R. Lowenstein, "Molecular indetermination in the transition to error catastrophe: systematic elimination of lymphocytic choriomeningitis virus through mutagenesis does not correlate linearly with large increases in mutant spectrum complexity," Proc. Natl. Acad. Sci. U. S. A. **99**, 12938–12943 (2002).

[27] W. E. Severson, C. S. Schmaljohn, A. Javadian, and C. B. Jonsson, "Ribavirin causes error catastrophe during Hantaan virus replication," J. Virology **77**, 481–488 (2003).

[28] S. Zhou, R. Liu, B. M. Baroudy, B. A. Malcolm, and G. R. Reyes, "The effect of ribavirin and IMPDH inhibitors on Hepatitis C virus subgenomic replicon RNA," Virology **310**, 333–342 (2003).

[29] J. A. Cuesta, "Huge progeny production during transient of a quasi-species model of viral infection, reproduction and mutation," Math. Comp. Model. **54**, 1676–1681 (2011).

[30] J. A. Cuesta, J. Aguirre, J. A. Capitán, and S. C. Manrubia, "Struggle for space: viral extinction through competition for cells," Phys. Rev. Lett. **106**, 028104–028104 (2011).

[31] J. A. Capitán, J. A. Cuesta, S. C. Manrubia, and J. Aguirre, "Severe hindrance of viral infection propagation in spatially extended hosts," PLoS One **6** (2011), 10.1371/journal.pone.0023358.

[32] H. W. Watson and F. Galton, "On the probability of the extinction of families," J. Anthropol. Inst. Great Britain

and Ireland **4**, 138–144 (1874).
[33] R. Miralles, P. J. Gerrish, A. Moya, and S. F. Elena, "Clonal interference and the evolution of rna viruses," Science **285**, 1745–1747 (1999).
[34] M. Imhof and C. Schlotterer, "Fitness effects of advantageous mutations in evolving escherichia coli populations," Proc. Natl. Acad. Sci. U. S. A. **98**, 1113–1117 (2001).
[35] P. D. Keightley and M. Lynch, "Toward a realistic model of mutations affecting fitness," Evolution **57**, 683–685; discussion 686–689 (2003).
[36] H. A. Orr, "The distribution of fitness effects among beneficial mutations," Genetics **163**, 1519–1526 (2003).
[37] P. Carrasco, F. de la Iglesia, and S. F. Elena, "Distribution of fitness and virulence effects caused by single-nucleotide substitutions in *Tobacco Etch Virus*," J. Virology **18**, 12979–12984 (2007).
[38] A. Eyre-Walker and P. D. Keightley, "The distribution of fitness effects of new mutations," Nat. Rev. Genet. **8**, 610–618 (2007).
[39] M. Parera, G. Fernàndez, B. Clotet, and M. A. Martínez, "Hiv-1 protease catalytic efficiency effects caused by random single amino acid substitutions," Mol. Biol. Evol. **24**, 382–387 (2007).
[40] D. R. Rokyta, C. J. Beisel, P. Joyce, M. T. Ferris, C. L. Burch, and H. A. Wichman, "Beneficial fitness effects are not exponential for two viruses," J. Mol. Evol. **67**, 368–376 (2008).
[41] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed., Vol. 1 (Wiley, New York, 1968).
[42] F. Gantmacher, *Applications of the Theory of Matrices* (Dover Publications, New York, 2005).
[43] C. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, Philadelphia, 2000).
[44] T.E. Harris, *The Theory of Branching Processes* (Springer-Verlag, Berlin, 1963).
[45] K. B. Athreya and P. E. Ney, *Branching Processes* (Springer-Verlag, Berlin, 1972).
[46] M. Kimmel and D.E. Axelrod, *Branching Processes in Biology* (Springer-Verlag, New York, 2002).
[47] P. Jagers, F. C. Klebaner, and S. Sagitov, "On the path to extinction," Proc. Natl. Acad. Sci. U. S. A. **104**, 6107–6111 (2007).
[48] M. Jiřina, "A simplified proof of the Sevastyanov theorem on branching processes," Annales de l'I. H. Poincaré, Section B **6**, 1–7 (1970).
[49] H. Kesten and B. P. Stigum, "A limit theorem for multidimensional Galton-Watson processes," Ann. Math. Stat. **37**, 1211–1223 (1966).
[50] H. Kesten and B. P. Stigum, "Additional limit theorems for indecomposable multidimensional Galton-Watson processes," Ann. Math. Stat. **37**, 1463–1481 (1966).
[51] H. Kesten and B. P. Stigum, "Limit theorems for decomposable multi-dimensional Galton-Watson processes," J. Math. Annal. Appl. **17**, 309–338 (1967).
[52] R. Lyons, R. Pemantle, and Y. Peres, "Conceptual proofs of $L \log L$ criteria for mean behaviour of branching processes," Ann. Probab. **23**, 1125–1138 (1995).
[53] P. Olofsson, "The $x \log x$ condition for general branching processes," J. Appl. Probab. **35**, 537–544 (1998).
[54] T. G. Kurtz, R. Lyons, R. Pemantle, and Y. Peres, "A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes," in *Classical and Modern Branching Processes*, IMA Vol. Math. Appl., Vol. 84, edited by K. Athreya and P. Jagers (Springer-Verlag, New York, 1994) pp. 181–185.
[55] T. W. Mullikin, "Limiting distributions for critical multitype branching processes with discrete time," Trans. Am. Math. Soc. **106**, 469–494 (1963).
[56] A. Joffe and F. Spitzer, "On multitype branching processes with $\rho \leqslant 1$," Annales de l'I. H. Poincaré, Section B **19**, 409–430 (1967).
[57] J. Foster and P. Ney, "Limit laws for decomposable critical branching processes," Z. Wahrscheinlichkeitstheorie verw. Gebiete **46**, 13–43 (1978).
[58] H. J. Muller, "The relation of recombination to mutational advance," Mutat. Res. **106**, 2–9 (1964).
[59] R. Forster, C. Adami, and C. O. Wilke, "Selection for mutational robustness in finite populations," J. Theor. Biol. **243**, 181–190 (2006).
[60] S. F. Elena, C. O. Wilke, C. Ofria, and R. E. Lenski, "Effects of population size and mutation rate on the evolution of mutational robustness," Evolution **61**, 666–674 (2007).
[61] P. Domingo-Calap, J. M. Cuevas, and R. Sanjuán, "The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages," PLoS Genet **5**, e1000742 (2009).
[62] M. H. Katz and J. L. Gerberding, "Postexposure treatment of people exposed to the human immunodeficiency virus through sexual contact or injection-drug use," N. Engl. J. Med. **336**, 1097–1100 (1997).
[63] M. Eigen, "Selforganization of matter and the evolution of biological macromolecules," Naturwissenschaften **58**, 465–523 (1971).
[64] I. Leuthäusser, "Statistical mechanics of Eigen's evolution model," J. Stat. Phys. **48**, 343–360 (1987).
[65] J. J. Bull, L. A. Meyers, and M. Lachmann, "Quasispecies made simple," PLoS Comp. Biol. **1**, 450–460 (2005).
[66] S. C. Manrubia, E. Domingo, and E. Lázaro, "Pathways to extinction: beyond the error threshold," Phil. Trans. R. Soc. B **365**, 1943–1952 (2010).
[67] O. Riordan and L. Warnke, "Explosive percolation is continuous," Science **333**, 322–324 (2011).
[68] R. M. Karp, "The transitive closure of a random digraph," Random Struct. Alg. **1**, 73–93 (1990).
[69] B. Bollobás and O. Riordan, "Random graphs and branching processes," in *Handbook of Large-Scale Random Networks*, Bolyai Society Mathematical Studies, Vol. 18, edited by B. Bollobás, R. Kozma, and D. Miklós (Springer-Verlag, 2008) Chap. 1, pp. 15–115.
[70] R. M. Ribeiro, L. Qin, L. L. Chavez, D. Li, S.G. Self, and A. S. Perelson, "Estimation of the initial viral growth rate and basic reproductive number during acute hiv-1 infection," J. Virology **84**, 6096–6102 (2010).
[71] D. Castro, *Simulação Computacional e Análise de um Modelo Fenotípico de Evolução Viral*, Master's thesis, Universidade Federal de São Paulo - UNIFESP, São Paulo (2011).
[72] D. Castro, F. Antoneli, F. Bosco, and L. M. Janini, In preparation (2011).